



CHES

Chronic Headache Education
and Self-management Study

STATISTICAL ANALYSIS PLAN

Version: 1.1

Contents

SECTION 1: ADMINISTRATIVE INFORMATION	3
SECTION 2: INTRODUCTION	5
SECTION 3: STUDY METHODS.....	7
SECTION 4: STATISTICAL PRINCIPLES	10
SECTION 5: TRIAL POPULATION	12
SECTION 6: ANALYSIS	17
SECTION 7: TEMPLATE TABLES AND FIGURES.....	24
REFERENCES	25

SECTION 1: ADMINISTRATIVE INFORMATION

SECTION 1: ADMINISTRATIVE INFORMATION

Title: *Chronic Headache and Self-management Study (CHESS)*

ISRCTN number: 79708100

SAP Version: *Version 1.1 (Date: 16 July 2019)*

Protocol Version: *Version 3.4 (Date: 12 July 2018)*

SAP revisions: None

Roles and responsibility:

- Dr Dipesh Mistry, Warwick Clinical Trials Unit (WCTU) – Trial Statistician (Author of SAP)
- Dr Siew Wan Hee, Warwick Medical School (WMS) – Statistician (Co-applicant)
- Professor Sandra Eldridge, Barts and The London School of Medicine and Dentistry – Senior Statistician (Co-applicant)
- Professor Martin Underwood, Warwick Clinical Trials Unit (WCTU) – Principal Investigator

Signatures of:

	Name	Date	Signature
Author of SAP	Dipesh Mistry		
Co-applicant	Siew Wan Hee		
Senior statistician	Sandra Eldridge		
Principal investigator	Martin Underwood		

SECTION 2: INTRODUCTION

SECTION 2: INTRODUCTION

Background and rationale

Chronic headaches present a major problem both for the individual and society. Previous studies on supportive self-management interventions in this population have largely been small studies with short term follow-up, they often did not report clinically relevant outcomes, or were conducted in different healthcare systems therefore difficult to translate into an NHS setting. These studies also did not necessarily focus on chronic headache but rather looked at headache with no frequency specified. Based on the results of our systematic review there may be potential for large gain through a combination of self-management education and appropriate use of prophylaxis and management of medication overuse headache in a chronic headache population.

In order to develop the evidence base needed for self-management intervention for chronic headache there needs to be a carefully developed, piloted and evaluated intervention package which has been supported by good qualitative work on understanding outcomes of interest. There is therefore the need for a robust clinical and cost-effectiveness trial within an NHS setting.

Objectives

The objective is to answer the question: Amongst adults with chronic headache arising from migraine, chronic tension type headache or medication overuse headache, is the provision of a self-management support programme in addition to best usual NHS care clinically and cost effective?

SECTION 3: STUDY METHODS

SECTION 3: STUDY METHODS

Trial design

This trial is a multi-centre randomised controlled trial comparing a group education and self-management intervention with a best usual care plus relaxation control for participants living with chronic tension type headaches, probable chronic migraine or definite chronic migraine with or without medication overuse headache.

Randomisation

The randomisation allocation ratio is 1:1.07 due to the method used to compute the sample size with clustering in one arm. Randomisation will be stratified by geographical locality (Midlands and Greater London) and headache type (six possible headache types; chronic tension type headache, probable chronic migraine and definite chronic migraine with or without medication overuse headache) using minimisation. Randomisation will take place using an online application specifically developed for the CHESS Study by the Warwick CTU programming team. (See section 2.6.3 of the protocol).

Sample size

A detailed description of the sample size calculation can be found in section 5.8 of the protocol. In brief, a sample size of 689 (333 in the relaxation arm and 356 in the self-management programme) will provide 90% power to detect a between group difference in those with migraine of 2 (SD: 6.9) in the HIT-6 score measured at 12 months at the two-sided 5% significance level. The sample size also accounted for 20% loss to follow-up and clustering in the self-management arm using an intra-class correlation coefficient (ICC) of 0.01 assuming an average group size of 10.

Framework

A superiority hypothesis testing framework will be used to compare the self-management arm to the relaxation arm.

Statistical interim analyses and stopping guidance

There are no planned interim analyses or stopping guidelines for this study. However, in consultation with the Data Monitoring Committee (DMC) we would review the sample size around halfway through recruitment to ensure we have sufficient participants with probable or definite chronic migraine. If the proportion of participants with chronic tension type headache is $\leq 15\%$ then we will recruit more participants with probable or definite chronic migraine such that we could perform the primary clinical analysis on this subpopulation.

Timing of final analysis

Once all of the data has been collected from participants, entered onto the database, fully validated and cleaned, the database will then be locked. The final analyses on all outcomes will then be conducted at each of the follow-up time points.

Timing of outcome assessments

Primary and secondary outcomes will be collected at baseline, 4, 8 and 12 months follow-up.

SECTION 4: STATISTICAL PRINCIPLES

SECTION 4: STATISTICAL PRINCIPLES

Confidence intervals and P values

All statistical tests will be two-sided at the 5% significance level. The estimate, 95% confidence interval (95% CI) and P value will be reported for each test undertaken.

Adherence and protocol deviations

We will look at two levels of adherence in this study; minimal adherence and full adherence. Minimal adherence with the intervention is defined as the participant attending day 1 of the intervention plus the one-to-one session. Full adherence is defined as the participant attending both days, plus individualised contact with the nurse. Both levels of adherence will inform the complier averaged causal effect (CACE) analysis.

Analysis populations

All analyses will be available case analyses based on 'Intention-to-treat' (ITT) principles. Participants will be analysed according to the treatment they were randomised to, irrespective of the treatment they actually received. All participants will be included in the analysis, regardless of whether they adhered to the protocol. The main summary tables and analyses will be based on the intention-to-treat population.

SECTION 5: TRIAL POPULATION

SECTION 5: TRIAL POPULATION

Screening data

A detailed summary of the screening data will be presented as frequencies and percentages to describe the representativeness of the trial sample. The screening summary will start at the GP practice population search level (i.e. how many practices were approached, the number records searched, the number of mail outs etc.) right the way through to final consent and randomisation. This will also include a summary of how many participants were self-referrals and how many were approached via the GP practice.

Eligibility

Patients are eligible to be included in the trial if they meet the following criteria:

Inclusion criteria

- Able and willing to comply with the study procedures and provision of written informed consent.
- Aged ≥ 18 years.
- Living with chronic headache; defined as headache on 15 or more days per month for at least three months.
- Result of nurse classification interview confirms headache type to be definite or probable chronic migraine, or chronic tension type headache, with or without medication overuse headache.
- Fluent in written and spoken English.

Exclusion criteria

- Unable to attend the group sessions.
- No access to a telephone.
- Has an underlying serious psychological disorder with ongoing symptoms which preclude or significantly interfere with participation in the group intervention.
- Previous entry or randomisation in the present trial.

- Is currently participating in another clinical trial of headache treatments, or in a trial of an unregistered medicinal product, or less than 90 days have passed since completing participation in such a trial.

The eligibility will be summarised using frequencies and percentages to describe how many people were:

- Eligible and randomised
- Eligible and not randomised
- Ineligible and randomised (in error)
- Ineligible and not randomised; summarising the main reasons for exclusion

In addition to the above, a summary of the different headache types identified from the nurse classification interviews will also be presented (definite or probable chronic migraine, or chronic tension type headache, with or without medication overuse).

Recruitment

The CONSORT diagram will illustrate the flow of participants throughout the trial. This will include:

- Number screened
- Of those screened, how many ineligible or declined
- Number randomised
- How many withdrew, died and were lost to follow-up at each follow-up time-point
- How many included in the final analyses at the primary endpoint listing reasons why participants were excluded

Withdrawal/follow-up

All withdrawals will be summarised by group using frequencies and percentages.

Level of withdrawal - will be summarised by treatment group i.e. how many withdrew from intervention alone but remained on follow-up and/or how many withdrew completely.

Timing of withdrawal – withdrawal timings in this trial will be summarised by treatment group as follows:

- Withdrawals after randomisation but before first group session (intervention arm only);
- Withdrawals during group sessions (intervention arm only);
- Withdrawals from follow-up - (i) withdrawal prior to 4-month follow-up (ii) withdrawal after 4-month follow-up but before 8-month follow-up (iii) withdrawal after 8-month follow-up but before 12-month follow-up

Withdrawal decision – the withdrawal decision i.e. decision made by participant or CHESS study team, will be summarised by treatment group

Withdrawal reason – participants have the option to provide a reason for withdrawal if they withdraw. Withdrawal reasons will be summarised.

Follow-up rates - follow-up rates are based on case report form (CRF) completion at follow-up time points.

$$\% \text{ Follow-up rate (at time T)} = \frac{\text{Number of participants assessed at time T}}{\text{Total no. that should have been assessed at time T}} \times 100$$

Follow-up rates will be computed at the 4-, 8- and 12-month follow-up time-points.

Baseline patient characteristics

The demographic characteristics and pre-randomisation clinical outcome measures of all randomised participants will be summarised by treatment allocation. The table below lists the demographic and clinical measures that will be collected.

Type of Data	Outcome measures
Demographic:	<ul style="list-style-type: none"> - Age - Gender - Ethnic group - Age at leaving full time education - Current work status

Clinical measures:	
General Health	<ul style="list-style-type: none"> - Fatigue - Sleep quality - Bodily pain - Troublesomeness grid
Headache Specific	<ul style="list-style-type: none"> - Headache Specific Information (HIT-6)[1] - Chronic Headache Quality of Life Questionnaire, version1.0 (CHQLQ) [2] - Headache frequency, severity and duration over the past 7 days
Health-related Quality of Life	<ul style="list-style-type: none"> - Short Form 12-item Health Survey (SF12 (v2)) [3] - EuroQoL (EQ5D-5L) [4] - Chronic Headache Quality of Life Questionnaire, version1.0 (CHQLQ) [2]
Mood	Hospital Anxiety and Depression Scale (HADS) [5]
Confidence	Pain Self-Efficacy Questionnaire (PSEQ) [6]
Social Activity	Social Integration Subscale (heiQ) [7]
Health economic measures:	
Medication	<ul style="list-style-type: none"> - Medication purchased in last four weeks over the counter - Cost
Healthcare Use	<ul style="list-style-type: none"> - Inpatient care - Admission details - NHS Day Care treatment - Community health and social care - Side effects from headache medication - Private treatment - Additional cost information

For continuous data, the number of participants (n), mean, standard deviation (SD), median and interquartile range (IQR) will be used to summarise the outcome measures by treatment allocation. The number (%) of participants will be used to summarise categorical outcome measures.

SECTION 6: ANALYSIS

SECTION 6: ANALYSIS

Outcome definitions

The table below lists and describes the primary and secondary outcomes. This includes details of specification of outcomes, timings and the derivation of the outcome (if required).

Outcome	Time point	Derivation of outcome
Primary outcome		
HIT-6 score[1]	1, 2, 3, 4	HIT-6 consists of 6 questions, each with 5 responses (never to always) which are scored 6, 8, 10, 11, and 13 points respectively. The HIT-6 is computed by simply summing the scores across the 6 questions. The score ranges from 36-78; the higher the score the greater the severity of headache.
Secondary outcomes		
Chronic Headache Quality of Life Questionnaire, version 1.0 (CHQLQ)	1, 2, 3, 4	Measures chronic headache quality of life on scale of 0-100 over 3 domains (role restrictive, role preventive and emotional function). A higher score indicates better quality of life.
SF-12 V2 [3]	1, 2, 3, 4	SF-12 score computed using the algorithm/software provided by the authors. The algorithm produces mental and physical component scores ranging from 0-100 where a higher score reflects better mental and physical functioning, respectively.
EQ-5D-5L [4]	1, 2, 3, 4	EQ-5D-5L score will be computed in Stata using the eq5d package. The EQ-5D-5L score ranges from 0-1 where a higher score reflects better quality of life.
Hospital Anxiety and Depression Scale (HADS) [5]	1, 2, 3, 4	The HADS consists of 14 questions each with 4 responses with an assigned score. Seven questions measure anxiety and the other seven measure depression. The scores are simply summed up to give an anxiety and depression score both ranging from 0-21 where a higher score reflects more severe anxiety and depression.
Pain Self-Efficacy Questionnaire (PSEQ) [6]	1, 2, 3, 4	PSEQ consists of 10 questions, each with 6 responses (Not at all confident to Completely confident) which are scored from 0-6 respectively. The PSEQ is computed by simply summing the scores across the 10 questions. The score ranges from 0-60 where higher scores reflect stronger self-efficacy beliefs.
Social Integration Subscale of the Health Education Impact Questionnaire (heiQ) [7]	1, 2, 3, 4	The Social Integration subscale of heiQ measures the impact of social engagement and support through interaction with others presented with the same illness. If >50% questions present then values can be assigned for scoring otherwise the score is missing. Higher scores indicate higher level of social interaction.

Headache days (Collected via smartphone app or paper version)	Once a week for the first 6 months and then once a month for the following 6 months. Also collected at 1, 2, 3, 4.	Data collected on: <ul style="list-style-type: none"> - On how many of the last 7 days have you had a headache - On those days, on average how long did they last - On those days, on average how severe were they
Safety reporting		
Adverse Events and Serious Adverse Events	Throughout the trial	

1 Baseline

2 4 month after randomisation

3 8 months after randomisation

4 12 months after randomisation

Analysis methods

Participant characteristics and outcomes will be summarised as mean and standard deviation (SD) for continuous data or frequency and percentage for categorical data, summarised by treatment arm. The median and interquartile range (IQR) will be presented if data are non-normal.

The primary analysis approach will be intention to treat. To account for the trial design with clustering in the intervention arm, linear mixed effects models with partial clustering will be used to estimate treatment effects for both primary and secondary outcomes. This will be done using the *mixed* command in Stata. Analyses will be adjusted for age, gender, the baseline value of the dependent variable and baseline stratification factors (type of headache and geographical locality). The adjusted treatment effect estimates (mean difference) will be presented along with their associated 95% confidence interval (CI). The primary clinical analysis will assess the overall difference between the self-management therapy (intervention) and the relaxation therapy (control) groups in the population with either probable or definite chronic migraine (if the proportion of participants with chronic tension type headache is $\leq 15\%$). If the proportion of chronic tension type headache is $>15\%$ then the primary analysis will be according to the whole population of chronic headache (chronic migraine and tension type headache).

The values of the variable “number of headache days in the past month” collected at baseline and each follow-up time point is in the range 0 to 28. As such a normal distribution may not be a suitable distribution to explain its frequency. We will therefore plot the frequency of headache days and explore whether other distributions, e.g. negative binomial and beta-binomial may be able to explain the data frequency. The plots will be examined visually before a distribution is assumed for the variable for further analysis. If more than one distribution is considered to be sufficient for the data then they will be used for further analyses and all the results will be presented. We may also explore the possibility of transforming the number of headache day’s data into proportion (or rates) or categorising the data into ordinal outcomes. The latter approach would decrease the precision and sensitivity of the outcome but may be better than assuming it follows an incorrect distribution.

The possibility of carrying out a complier averaged causal effect (CACE) analysis for the primary outcome will be explored. Pre-specified subgroup analyses will also be conducted using formal statistical tests for interaction to examine whether baseline anxiety, depression and severity are moderators of treatment effect.[8]

Missing data

The levels and patterns of non-responders at each follow-up time point (including the weekly/monthly headache days collected via the smartphone app) will be monitored regularly. This is to ensure that strategies could be identified and implemented to minimise non-responders.

The levels and patterns of missingness in the primary outcome will be assessed to determine the type of missingness (e.g. MAR, NMAR). If required, as an additional sensitivity analysis, imputation techniques relevant to the type of missing data mechanism will be used to impute data and estimate the treatment effect to see how it compares to the main ITT analysis.

Additional analyses

In addition to the primary analyses, the overall result for those with all headache types will also be assessed. NICE was specifically interested in data on specific headache types; rejecting data that reported data on a mixed population of people with chronic headaches. Therefore in addition to the primary analyses, the results (mean difference and 95% CI) for each of the three headache types separately, and the results for those with or without medication overuse separately will also be presented to facilitate future meta-analyses and inform future condition specific guidelines.

Data on total headache days was collected from participants over the whole study period. Participants had a choice of reporting this outcome either using a smartphone app or diary records (not both). This data was also collected in the baseline and follow-up questionnaires. We will compare the total headache days between the two groups using an area under the curve (AUC) approach. If participants have reported headache days data using both the app/diary and the follow-up form at the same time point, then we will use the app/diary as the primary data source. We expect there will be missing data. Therefore we will apply the following scoring algorithm in order to obtain a complete set of scores for each participant thus allowing us to undertake the AUC analysis.

- A blank score for each expected observation is created.
- If there is a valid text message response for the expected observation, then the blank value is replaced with the text message score.
- If the participant did not register with the text messaging service and there is a valid paper diary score, then the blank value is replaced with this paper diary score.
- If the participant did not provide a score via either the text messaging service or the diary, but a valid score is available on the follow-up form, then the blank value is replaced with the follow-up form score.
- If the participant has completed only one data source (either text message or paper diaries) and observation X is missing in the middle of the data set, then the score for observation X is calculated as:

$$\frac{(Obs\ X - 1) + (Obs\ X + 1)}{2} \quad (1)$$

- If two or more adjacent scores are missing, then a monotonic assumption is made for the missing values between the most recent valid score and the next available valid score. For example if two consecutive scores are missing, observation X and observation $X + 1$, then the scores for observation $X - 1$ and observation $X + 2$ are used to calculate the imputed values for observations X and day $X + 1$ as follows:

$$Obs\ X = Obs\ X - 1 + \frac{(Obs\ X + 2) - (Obs\ X - 1)}{Number\ of\ missing\ obs + 1} \quad (2)$$

$$Obs\ X + 1 = Obs\ X + \frac{(Obs\ X + 2) - (Obs\ X - 1)}{Number\ of\ missing\ obs + 1} \quad (3)$$

- If the participant has provided a score via both app/diary and the follow-up data, then the app/diary data is used.
- If the participant has complete both data sources but the app/diary score is missing, then the follow-up data score is used.
- If the first observation is missing then the first valid observation for this participant is backfilled.
- If last expected observation is missing, then the score from the 12 month follow-up will be used. If this score is missing then the last observation carried forward is used.

Around 30 participants will be included in the process evaluation interviews conducted from pre-randomisation to follow-up. It is possible that discussing their expectations before and during the study may influence the treatment effectiveness. A sensitivity analysis will therefore be performed that excludes these participants from the main analysis.

At the eligibility check, participants are eligible if they have chronic headache defined as 15 or more days of headache per month for at least three months. However on the baseline form, participants are asked to report the number of headache days over the last 4 weeks for which many report having less than 15 days of headache. A sensitivity analysis will therefore be performed that excludes these participants from the main analysis.

Harms

The frequency and percentage (%) of serious adverse events (SAE) and adverse events (AE) in the trial will be compared between the two treatments using the chi-squared test provided the expected values in the cross-tabulation are greater than five, otherwise Fisher's exact test will be used. Odds ratios and 95% confidence intervals will be reported. Adjusted analyses will not be performed for any harm data. The event type, severity assessment, expectedness and relatedness to intervention will also be summarised by treatment arm.

Statistical software

Statistical analyses will be conducted using the statistical software package Stata 15.0.

SECTION 7: TEMPLATE TABLES AND FIGURES

SECTION 7: TEMPLATE TABLES AND FIGURES

The template tables and figures have been presented in a separate document that consists of the following sections:

SECTION 1 - Screening through to randomisation

SECTION 2 - Participant baseline and demographic data

SECTION 3 - Participant follow-up

SECTION 4 - Intervention data

SECTION 5 - Study outcome data

SECTION 6 - Adverse events and serious adverse events

REFERENCES

- [1] Kosinski, M., et al., *A six-item short-form survey for measuring headache impact: the HIT-6*. Qual Life Res, 2003. **12**(8): p. 963-74.
- [2] McKenna, S.P., L.C. Doward, and K.M. Davey, *The Development and Psychometric Properties of the MSQOL: A Migraine-Specific Quality-of-Life Instrument*. Clin Drug Investig, 1998. **15**(5): p. 413-23.
- [3] Jenkinson, C. and R. Layte, *Development and Testing of the UK SF-12*. Journal of Health Services Research, 1997. **2**(1): p. 14-18.
- [4] Herdman, M., et al., *Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L)*. Qual Life Res, 2011. **20**(10): p. 1727-36.
- [5] Zigmond, A.S. and R.P. Snaith, *The hospital anxiety and depression scale*. Acta Psychiatr Scand, 1983. **67**(6): p. 361-70.
- [6] Nicholas, M.K., *The pain self-efficacy questionnaire: Taking pain into account*. Eur J Pain, 2007. **11**(2): p. 153-63.
- [7] Osborne, R.H., G.R. Elsworth, and K. Whitfield, *The Health Education Impact Questionnaire (heiQ): an outcomes and evaluation measure for patient education and self-management interventions for people with chronic conditions*. Patient Educ Couns, 2007. **66**(2): p. 192-201.
- [8] Brookes, S.T., et al., *Subgroup analyses in randomised controlled trials: quantifying the risks of false-positives and false-negatives*. Health Technol Assess, 2001. **5**(33): p. 1-56.